PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Smartphone problem-solving and behavioral activation therapy to
	reduce fear of recurrence among breast cancer patients
	(SMartphone Intervention to LEssen fear of cancer recurrence:
	SMILE project): protocol for a randomized controlled trial
AUTHORS	Akechi, Tatsuo; Yamaguchi, Takuhiro; Uchida, Megumi; Imai, Fuminobu; Momino, Kanae; Katsuki, Fujika; Sakurai, Naomi; Miyaji, Tempei; Horikoshi, Masaru; Furukawa, Toshi; Iwata, Hiroji; Uchitomi, Yosuke

VERSION 1 – REVIEW

REVIEWER	Aaron Heller
	University of Miami/USA
REVIEW RETURNED	25-Jul-2018
GENERAL COMMENTS	Akechi and colleagues present an interesting trial using smartphones to reduce anxiety and fear of recurrence in women who have survived breast cancer. They have already suggested at the effectiveness of an in-person intervention (a combination of Problem Solving and Behavioral Activation) to reduce fear of cancer recurrence. They are attempting to increase the reach of this intervention via mobile technology. This is an interesting idea and well worth pursuing, with one major flaw of the study:
	The authors propose only two groups: a wait-list control vs. the smartphone treatment group. Why do the authors not propose a third group, which is the in-person PST/BA treatment? They have already demonstrated that the in-person approach has some effectiveness, and being able to compare effect sizes in this group to the other two groups would greatly enhance the impact of this study. Second, I would encourage the authors to collect a measure of guality of life, specifically for breast cancer such as the FACT-B.

REVIEWER	Courtney Beard
	McLean Hospital/Harvard Medical School United States of America
REVIEW RETURNED	09-Aug-2018
GENERAL COMMENTS	p. 7 – the pilot data for the BA app reportedly included cognitive restructuring. It is not clear whether or not the cognitive restructuring component of the app is included in the current trial.
	Introduction – authors provide preliminary evidence of efficacy for PST and BA, but no rationale for what mechanisms these interventions are targeting in FCR. BA is typically used to treat

might be an additional interesting measure to include.

 depression, and it is not immediately clear why solving problems would help fear of recurrence. p. 9 - the authors provide substantially less detail about the BA sessions than the PST sessions. What exactly are participants doing or being taught in the BA sessions? Currently, it is impossible to know if it is actually BA being delivered. Thus, #11a on the checklist is not currently complete. p. 10 - it was surprising that eligibility criteria did not include fear of recurrence. It is not acceptable to have an exclusion criterion as vague as "judged inappropriate for participation by the researchers" - this would allow the researchers to pick and choose whoever they want in the study. There need to be specific reasons named. p. 15 - there is no description of treatment as usual. How will this be characterized or controlled in analyses? p. 15 - similarly - discontinuing a participant because "the research team judges that it is inappropriate to controue for any reason" is too vague and allows too much potential bias. p. 19 - consider using a different term than "effective parts". Asking participants about their evaluation of the app will not actually tell you which components are responsible for symptom improvement. p. 21 - 5% drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study map be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2). Figure 2 - it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session. Although the authors reportedly included the following information (according to their checklist), I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not description of participant compensati	
 p. 9 – the authors provide substantially less detail about the BA sessions than the PST sessions. What exactly are participants doing or being taught in the BA sessions? Currently, it is impossible to know if it is actually BA being delivered. Thus, #11a on the checklist is not currently complete. p. 10 – it was surprising that eligibility criteria did not include fear of recurrence. It is not acceptable to have an exclusion criterion as vague as "judged inappropriate for participation by the researchers" – this would allow the researchers to pick and choose whoever they want in the study. There need to be specific reasons named. p. 15 – there is no description of treatment as usual. How will this be characterized or controlled in analyses? p. 15 – similarly – discontinuing a participant because "the research team judges that it is inappropriate for symptom improvement. p. 19 – consider using a different term than "effective parts". Asking participants about their evaluation of the app will not acclually tell you which components are responsible for symptom improvement. p. 21 – 5% drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study may be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2). Figure 2 – it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session. Atthough the authors reported to could not find this. Authors report that p.8 includes explanation of choice of comparators -1 don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration continue for any term of a data monitoring committee. # 22 on checklist – did not see anything on plans for collecting information ac	depression, and it is not immediately clear why solving problems would help fear of recurrence.
 p. 10 - it was surprising that eligibility criteria did not include fear of recurrence. It is not acceptable to have an exclusion criterion as vague as "judged inappropriate for participation by the researchers" - this would allow the researchers to pick and choose whoever they want in the study. There need to be specific reasons named. p. 15 - there is no description of treatment as usual. How will this be characterized or controlled in analyses? p. 15 - similarly - discontinuing a participant because "the research team judges that it is inappropriate to continue for any reason" is too vague and allows too much potential bias. p. 19 - consider using a different term than "effective parts". Asking participants about their evaluation of the app will not actually tell you which components are responsible for symptom improvement. p. 21 - 5% drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study may be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2). Figure 2 - it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session. Although the authors reportedly included the following: No description of participant compensation. p. 24 - authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators - 1 don't see this. Not clear number of and type of sites Discontinuation or circleria do not include anything about clinical deterioration. Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. # 32 on checklist - no model consent form provided	p. 9 – the authors provide substantially less detail about the BA sessions than the PST sessions. What exactly are participants doing or being taught in the BA sessions? Currently, it is impossible to know if it is actually BA being delivered. Thus, #11a on the checklist is not currently complete.
 p. 15 – there is no description of treatment as usual. How will this be characterized or controlled in analyses? p. 15 – similarly – discontinuing a participant because "the research team judges that it is inappropriate to continue for any reason" is too vague and allows too much potential bias. p. 19 – consider using a different term than "effective parts". Asking participants about their evaluation of the app will not actually tell you which components are responsible for symptom improvement. p. 21 – 5% drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study may be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2). Figure 2 – it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session. Although the authors reportedly included the following information (according to their checklist), I could not find the following: No description of participant compensation. p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. # 32 on checklist – no model consent form provided 	 p. 10 – it was surprising that eligibility criteria did not include fear of recurrence. It is not acceptable to have an exclusion criterion as vague as "judged inappropriate for participation by the researchers" – this would allow the researchers to pick and choose whoever they want in the study. There need to be specific reasons named.
 p. 15 - similarly – discontinuing a participant because "the research team judges that it is inappropriate to continue for any reason" is too vague and allows too much potential bias. p. 19 - consider using a different term than "effective parts". Asking participants about their evaluation of the app will not actually tell you which components are responsible for symptom improvement. p. 21 - 5% drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study may be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2). Figure 2 - it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session. Although the authors reportedly included the following information (according to their checklist), I could not find the following: No description of participant compensation. p. 24 - authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators - I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration. Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. # 32 on checklist - no model consent form provided 	p. 15 – there is no description of treatment as usual. How will this be characterized or controlled in analyses?
 p. 19 – consider using a different term than "effective parts". Asking participants about their evaluation of the app will not actually tell you which components are responsible for symptom improvement. p. 21 – 5% drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study may be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2). Figure 2 – it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session. Although the authors reportedly included the following: No description of participant compensation. p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. # 32 on checklist – no model consent form provided 	p. 15 – similarly – discontinuing a participant because "the research team judges that it is inappropriate to continue for any reason" is too vague and allows too much potential bias.
 p. 21 – 5% drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study may be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2). Figure 2 – it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session. Although the authors reportedly included the following information (according to their checklist), I could not find the following: No description of participant compensation. p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. #22 on checklist – did not see anything on plans for collecting information about harms. # 32 on checklist – no model consent form provided 	p. 19 – consider using a different term than "effective parts". Asking participants about their evaluation of the app will not actually tell you which components are responsible for symptom improvement.
 Figure 2 – it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session. Although the authors reportedly included the following information (according to their checklist), I could not find the following: No description of participant compensation. p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. #22 on checklist – did not see anything on plans for collecting information about harms. # 32 on checklist – no model consent form provided 	p. $21 - 5\%$ drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study may be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2).
 Although the authors reportedly included the following information (according to their checklist), I could not find the following: No description of participant compensation. p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. #22 on checklist – did not see anything on plans for collecting information about harms. # 32 on checklist – no model consent form provided 	Figure 2 – it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session.
No description of participant compensation. p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. #22 on checklist – did not see anything on plans for collecting information about harms. # 32 on checklist – no model consent form provided	Although the authors reportedly included the following information (according to their checklist), I could not find the following:
 p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. #22 on checklist – did not see anything on plans for collecting information about harms. # 32 on checklist – no model consent form provided 	No description of participant compensation.
Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee. #22 on checklist – did not see anything on plans for collecting information about harms. # 32 on checklist – no model consent form provided	 p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this. Authors report that p. 8 includes explanation of choice of comparators – I don't see this. Not clear number of and type of sites Discontinuation criteria do not include anything about clinical deterioration
 #22 on checklist – did not see anything on plans for collecting information about harms. # 32 on checklist – no model consent form provided 	Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee.
	 #22 on checklist – did not see anything on plans for collecting information about harms. # 32 on checklist – no model consent form provided

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 Reviewer Name: Aaron Heller Institution and Country: University of Miami/USA

Please leave your comments for the authors below

1) Akechi and colleagues present an interesting trial using smartphones to reduce anxiety and fear of recurrence in women who have survived breast cancer. They have already suggested at the effectiveness of an in-person intervention (a combination of Problem Solving and Behavioral Activation) to reduce fear of cancer recurrence. They are attempting to increase the reach of this intervention via mobile technology. This is an interesting idea and well worth pursuing, with one major flaw of the study:

Thank you for Dr. Heller's supportive and helpful comments.

2) The authors propose only two groups: a wait-list control vs. the smartphone treatment group. Why do the authors not propose a third group, which is the in-person PST/BA treatment? They have already demonstrated that the in-person approach has some effectiveness, and being able to compare effect sizes in this group to the other two groups would greatly enhance the impact of this study.

Thank you for your insightful comment. We agree with the significance of the third group as suggested. On the other hand, we decide this design comparing just two groups because of limited number of clinical research coordinators as well as limited financial resource.

3) Second, I would encourage the authors to collect a measure of quality of life, specifically for breast cancer such as the FACT-B, might be an additional interesting measure to include.

Thank you for your suggestive comment. As the reviewer pointed out, initial protocol has also included an outcome with regard to cancer patient's quality of life. However, after our research groups' careful discussion, quality of life measurements was finally excluded for some reasons. First reason is that to the best of our knowledge this is the first trial applies ePRO system via smartphone to obtain the patient's outcome and we would like to lessen patient's burden as possible as we can. Second reason is that our primary aim is improving patient's psychological function, not quality of life including physical function based on our need's survey (Akechi T et al, Patient's perceived need and psychological distress and/or quality of life in ambulatory breast cancer patients in Japan. Psychooncology. 2011 May;20(5):497-505).

Reviewer: 2

Reviewer Name: Courtney Beard

Institution and Country: McLean Hospital/Harvard Medical School, United States of America Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors

1) below p. 7 – the pilot data for the BA app reportedly included cognitive restructuring. It is not clear whether or not the cognitive restructuring component of the app is included in the current trial.

The smartphone CBT app that we have developed is an app for treatment of depression and is different from the apps used in the current trial.

2) Introduction – authors provide preliminary evidence of efficacy for PST and BA, but no rationale for

what mechanisms these interventions are targeting in FCR. BA is typically used to treat depression, and it is not immediately clear why solving problems would help fear of recurrence.

Thank for your helpful comment. We have added our hypothesis about how PST and BA contribute to improving FCR in the introduction section (p. 7, lines 8-12).

3) p. 9 – the authors provide substantially less detail about the BA sessions than the PST sessions. What exactly are participants doing or being taught in the BA sessions? Currently, it is impossible to know if it is actually BA being delivered. Thus, #11a on the checklist is not currently complete.

Thank for your helpful comment. We have added the explanation regarding the BA sessions (p. 10, lines 2-7).

4) p. 10 – it was surprising that eligibility criteria did not include fear of recurrence. It is not acceptable to have an exclusion criterion as vague as "judged inappropriate for participation by the researchers" – this would allow the researchers to pick and choose whoever they want in the study. There need to be specific reasons named.

As you know, there is no standard cut-off point of fear of recurrence. In addition, since we think that study subject focusing on patients who would like to participate in the intervention is practically meaningful, we did not include the level of fear of recurrence as the eligibility criteria. We agree with the inappropriate description of the suggested exclusion criteria. Since this is the first study using a newly developed research management system making full use of ICT technology, our discussion concluded that some possible problems, including identity theft, duplicate entry, and so on can occur. We have added these examples (p. 11, line 6).

5) p. 15 – there is no description of treatment as usual. How will this be characterized or controlled in analyses?

Since treatment as usual (TAU) means general treatment and/or care commonly provided by each patient's hospital (e.g., nurse's support etc.), we have added these explanation (p. 9, lines 1-2). So we will not control TAU factors in analyses.

6) p. 15 – similarly – discontinuing a participant because "the research team judges that it is inappropriate to continue for any reason" is too vague and allows too much potential bias.

We set this criteria for when some problems, including identity theft, duplicate entry, and so on is detected. We have added these examples (p. 15, line 16).

7) p. 19 – consider using a different term than "effective parts". Asking participants about their evaluation of the app will not actually tell you which components are responsible for symptom improvement.

We agree with the inappropriateness of the term. We changed to the term "perceived usability and/or merit" (p. 19, line 7).

8) p. 21 - 5% drop-out rate seems extremely optimistic. Is this based on the pilot studies? The study may be significantly underpowered if this estimate is wrong. Related, the pilot study effect appears tiny and not clinically meaningful (CARS-J score decrease from 12.8 to 11.2).

Thank you for your insightful comment. Drop-out rate is estimated by our pilot study investigating the effectiveness of Kaiketsu-App for 38 breast cancer survivors. The pilot study has shown that

completion rate of the study was 97.4% (Imai F, Akechi T et al, under review). As you know, there is no clear data regarding minimal clinically important difference of fear of recurrence. We agree that effect size of our Apps treatment suggests minimal to moderate effect of the intervention, however, we think that this power of the treatment can be still meaningful considering brevity of apps intervention.

9) Figure 2 – it is odd that the character who asked about PST is then the one explaining PST? It would be helpful to have an appendix for PST and BA that details what happens in each session.

Since one scene in a story is shown in Figure 2, talk between characters seems to be odd. We have changed to other figure (Figure 2). In addition, we describe details of PST and BA sessions in an appendix table as pointed out.

Although the authors reportedly included the following information (according to their checklist), I could not find the following:

We are very sorry for our confusing and/or incomplete reports. We have checked again and these are as follows:

10) No description of participant compensation.

We have added the information with regard to compensation (p. 20, lines 8-10).

11) p. 24 – authors refer to a description of checking adherence in the Methods and Analysis section, but I could not find this.

We have clearly added the information on adherence (p. 10, lines 9-11).

12) Authors report that p. 8 includes explanation of choice of comparators - I don't see this.

We have clearly added the information on rationale of comparator (p. 8, lines 9-11).

13) Not clear number of and type of sites

We have added the description as suggested (p. 11, lines 13-15).

14) Discontinuation criteria do not include anything about clinical deterioration

We have added the description as suggested (p. 15, lines 14-15).

15) Checklist reports that data monitoring description is on p. 12, but p. 12 is a table. I did not see any mention of a data monitoring committee.

We are very sorry for careless mistake. Since the psychological intervention provided by apps will not be invasive and also not produce serious harms, data monitoring committee will not be organized. We have added this explanation (p. 13, line 18-p.14, line 6).

16) #22 on checklist – did not see anything on plans for collecting information about harms.

No specific and serious adverse events are presumed in participants who use the Kaiketsu and Genki-Apps. However, using these apps might lead to psychological distress in some participants depending on their psychological state. We will evaluate these potential adverse events by qualitative

interview. We have added this description (p. 20, lines 2-6).

17) # 32 on checklist - no model consent form provided

We have provided our informed consent material on study's Web site (https://smile-project.org/).

VERSION 2 – REVIEW

REVIEWER	Aaron Heller
	University of Miami, USA
REVIEW RETURNED	05-Sep-2018

GENERAL COMMENTS	I understand that stretched resources will encumber the project's capacity to have a third arm of treatment as usual. However, I think the authors should note the lack of an active control arm as a limitation in the introduction or methods to dissect the specific mechanisms of change (and presumably the discussion once the data have been analyzed).

REVIEWER	Courtney Beard McLean Hospital
REVIEW RETURNED	17-Sep-2018

GENERAL COMMENTS	the authors have provided the missing details.
	one thing remains unclear. i would recommend describing what the
	"actual training" is and "getting the knack of behavioral activation."
	As a clinician who delivers behavioral activation and a researcher, I
	have no idea what this means.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1 Reviewer Name: Aaron Heller Institution and Country: University of Miami, USA

I understand that stretched resources will encumber the project's capacity to have a third arm of treatment as usual. However, I think the authors should note the lack of an active control arm as a limitation in the introduction or methods to dissect the specific mechanisms of change (and presumably the discussion once the data have been analyzed).

Thank you for your insightful comment and we agree with the comment. As we describe some limitations in the discussion section, we have added the suggested comment as one of the limitation of the current study (p. 26, lines 4-7).

Reviewer: 2 Reviewer Name: Courtney Beard Institution and Country: McLean Hospital

The authors have provided the missing details. One thing remains unclear. I would recommend describing what the "actual training" is and "getting the knack of behavioral activation." As a clinician who delivers behavioral activation and a researcher, I have no idea what this means.

Actual training is similar to so called "homework" in the cognitive-behavioral therapy. Genki-App includes a self-learning sheet for planning and doing pleasurable and new activity, and for evaluating achievement after conducting its activity. In addition, Genki-App provides 4 knack of behavioral activation (1. Start an activity to be able to conduct by yourself, 2. Divide big aim into some smaller ones, 3. Plan a schedule to conduct an activity, 4. Image a situation when you can do it well). We have added these detailed explanation in the method section (p. 10 line 4-8).